



King's Research Portal

DOI:

[10.1016/j.lingua.2017.05.007](https://doi.org/10.1016/j.lingua.2017.05.007)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Ji, Y., & Hohenstein, J. (2017). Conceptualising voluntary motion events beyond language use: A comparison of English and Chinese speakers' similarity judgments. *LINGUA*, 195, 57-71.

<https://doi.org/10.1016/j.lingua.2017.05.007>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

**Conceptualising voluntary motion events beyond language use: a comparison of
English and Chinese speakers' similarity judgments**

¹Yinglin Ji, ²Jill Hohenstein

¹Shenzhen University, PR China

²King's College London, UK

Corresponding author:

Yinglin Ji, Research Centre for Language and Cognition, Shenzhen University,

No. 3688, Nanhai Avenue, Shenzhen City 518060, PR China

Tel: 00 86 755 26749807

Email: yinglin.ji@szu.edu.cn

Abstract

This study explores the relationship between language and cognition by testing how monolingual speakers (3-year-olds, 8-year-olds and adults) who speak languages with partial typological differences in motion description (English and Chinese) respond to visual motion event stimuli in a triads matching task. The results suggest, first of all, that the two groups of 3-year-olds are predominantly path-oriented, as evidenced by their significantly longer fixation on path-match videos rather than manner-match videos in a preferential-looking scheme.

Using categorical measurement of overt choices, the older children and adults showed no particular preference for either manner- or path-similarity criteria. However, the analysis using continuous measurement of reaction time revealed that adults were significantly quicker in responding to voluntary motion stimuli than the older children of 8 years due to their cognitive maturity; manner-matched choices were decided in a significantly shorter time than path-matched choices, and Chinese participants were found to be significantly faster than their English peers in response. Overall, our results indicate that children's non-linguistic thought is similar prior to the internalisation of the lexicalisation patterns of motion events in their native languages, but it does not show language-specific divergences after such habitual use.

Keywords: motion event cognition, voluntary motion events, linguistic relativity hypothesis, triads matching task, preferential looking scheme

1. Introduction

The relationship between language, space and mind has been researched intensely in the last several decades. L. Talmy's work on the semantic framework of spatial expressions, the regular pattern of correspondence between motion components and grammatical categories in particular (Talmy 2000) laid the groundwork for an increasing number of significant findings in this field. Apart from the much-debated lexicalisation pattern in motion descriptions (i.e. satellite- vs. verb-framed), the study of the spatial domain touches upon the long-held perplexing question of the relationship between language and thought, relating to the Whorfian hypothesis.

With respect to motion event typology, crosslinguistic studies of motion event description have documented pervasive differences between languages. Much attention is currently devoted to the typological system *per se*: viz., is it more like a two (three)-way classification or a continuum, taking into account inter-category similarities as well as intra-category variations (Allen *et al.*, 2007; Bowerman and Choi, 2003; Chen and Guo, 2009; Hickmann, 2006; Ji *et al.*, 2011a, 2011b; Slobin, 1996, 2004)? The exact typological status of some languages (such as Chinese) is intensely debated and, more generally, the issue of typological variability within a given language is further explored (see, for instance, Ji *et al.* 2011a and Talmy 2009 for a discussion of Chinese, and Kopecka 2006 for a discussion of French). Such systematic language differences in motion expression raise fundamental questions for motion event cognition. If language indeed 'filters' or 'channels' the information we receive from the outside and thus constrains the way we perceive reality (i.e. as suggested by the Whorfian hypothesis), would speakers of different language groups conceptualise motion events differently?

Within this context, a growing number of researchers are looking at whether, and how, the effects of motion event typology can go beyond language use into a deeper level of mental conceptualisation. The current study aims to contribute to this 'cognitive' line of research by examining potential language influences on motion event cognition in children. Specifically, a non-verbal triads matching task has been used to assess the mental conceptualisation of voluntary motion events in Chinese and English participants who are aged 3, 8 and adult. The focus of the investigation is whether children's judgment of similarity in motion scenes is alike prior to acquiring the spatial language patterns exhibited by adults but shows differences after such acquisitions.

2. Motion representations at linguistic and cognitive levels

Given that our understanding of space is based on a universal image schema that has a kinaesthetic basis, spatial representation is traditionally thought to be universal. However, recent studies reveal that the linguistic representation of space shows striking typological differences across speakers of different language groups.

In Talmy's (2000) typological framework, some languages can be clearly classified as satellite-framed, in encoding manner of motion in verb roots and path in particles (English: A man *runs across* the road). Others are verb-framed, in expressing path of motion in verb roots and manner in adverbials (French: *Un homme traverse la route en courant* 'A man crosses the road by running'). However, some languages seem harder to 'fit' into these two categories. Thus, questions have been raised about serial verb languages (e.g. Chinese), which have been traditionally classified as satellite-framed, but are more recently considered to be a new, third type of 'equipollent-framed' language (Slobin 2004).

The influence of language-specific properties is found in various aspects of linguistic representation of space, which range from the coordinate system in spatial reference to the perception and expression of varied types of motion events (Berman and Slobin, 1994; Bowerman, 1999; Bowerman and Choi, 2003; Hickmann, 2006; Hickmann and Hendriks, 2010; Slobin, 1996, 2004, to name a few). It is also revealed to be present at different levels of linguistic analysis (e.g. morphology, lexicon, syntax and discourse) and across varied types of tasks (e.g., Allen *et al.*, 2007; Chen and Guo, 2009; Choi and Bowerman, 1991; Filipovic and Jaszczolt, 2012). To cite an example, Ji *et al.* (2011b) compared children's oral expression of motion events, in English and Chinese, in an elicited language production task. They examined in detail whether, and how, language-specific differences affect the course of children's acquisition of spatial language.

Their findings revealed that the utterance density was significantly higher for Chinese children than for age-matched English children. This is largely due to the fact that motion events are typically expressed in Chinese in a verb compound, which greatly facilitates the simultaneous encoding of multiple information components. Their investigation suggested, in particular, two 'critical periods' in children's acquisition of spatial language. Young children around the age of three years were found to be able to use typical motion expressions in their native language, though

their performance remained unstable and infrequent. However, by the age of 8, children's performance was reported to resemble that of adults fully, as evidenced by the number of motion components they encoded into their utterances and their ability to make finer distinctions regarding manner of motion (e.g., *walking*, *jogging*, *strolling*, *sauntering*) and to aggregate path details in a single clause (e.g., *climb up into the tree hole above from the ground*).

As for the effects of language typology on spatial cognition, the universal hypothesis stands in stark contrast to the view of linguistic relativity. The former holds that there is some cognitive universal shared by speakers of different languages, and any variations in surface linguistic encoding of motion events are no more than different instantiations of a single common conceptual framework (see, for instance, Jackendoff, 1996; Landau and Lakusta, 2006; McWhorter, 2014; Malt *et al.*, 2003). The latter, however, stresses that the effect of language typology extends beyond language use. Learning a language means acquiring a particular way of conceptualisation, and any differences in the mode of spatial cognition should be reflected, among other things, in language-related behaviour across speakers of different languages (see, for instance, Berthele, 2013; Boroditsky, 2000; Hohenstein, 2005; Levinson, 2003; Lucy, 1992; Zlatev and Blomberg, 2015).

Previous studies following the aforementioned views produced 'mixed' results. Some scholars reported no effect of language typology on cognition at all. For instance, Papafragou *et al.* (2002) compared the performance of English and Greek (satellite- vs. verb-framed) children and adults in memory and categorisation of motion events and found only similarity in speakers' behavioural responses. Other researchers employed more nuanced psychological schemes, such as preferential looking or eye tracking, to study how children across languages responded to visual motion signals (see, for instance, Dittmar *et al.*, 2008; Flecken *et al.*, 2015; Papafragou *et al.*, 2008; Von Stutterheim *et al.*, 2012). To illustrate, Hohenstein (2005) examined how Spanish- and English-speaking children responded to visual motion stimuli. She found that in a triads matching task, participants of different languages (verb-framed Spanish with 'path salience' vs. satellite-framed English with 'manner salience') behaved differently towards video stimuli in ways that could be predicted by their respective languages: 7-year-old English-speaking children fixated on videos matching the manner (rather than path) of a target video more often than Spanish-speaking 7-year-olds. Some other investigators detected certain forms of

language effects, but only under conditions in which language had been recruited for environmental signals, such as when participants were asked to verbally describe motion events just before they started performing non-verbal tasks (Dittmar *et al.*, 2008; Gennari *et al.*, 2002; Papafragou *et al.*, 2008; Soroli and Hickmann, 2010).

Needless to say, the great discrepancy in the findings of studies as reviewed above can be attributed to a variety of factors, amongst which the presence or absence of any verbal interference task seems to be particularly important. In Hohenstein's (2005) study, no verbal interference task was provided and, therefore, the significant differences in patterns of motion cognition between English and Spanish can be interpreted, in some sense, as reflecting language differences because it is possible that participants were subconsciously verbalising the video stimuli whilst watching and judging the similarity between events.

3. Linguistic encoding of motion events in English and Chinese

Motion events in English are characteristically encoded in a 'manner verb + path satellite' combination and the status of English as a representative satellite-framed language has been widely accepted. In contrast, the status of Chinese in motion event typology remains a controversial topic. Much of the debate centres around the exact grammatical category of the second constituent verb in a Chinese RVC (Resultative Verb Compound), which typically consists of, in sequence, a manner verb (V1), a path verb (V2) and an optional deictic verb (V3, e.g., *fei1-chu1-lai2* 'fly-exit/out-towards the speaker'). Talmy (1985, 2000, 2009) holds that V2 in an RVC is a verb-supporting element, which can be put on a par with English verb particles. One reason for this claim is that V2 in an RVC forms a closed-class set with a rather restricted number of members, just like verb particles. Furthermore, there is a semantic correspondence between English verb particles and the V2 in an RVC, both denoting the trajectory of motion.

Doubts have been raised regarding this classification. For example, Slobin (2004) found that Chinese actually possesses some clear verb-framing properties. The V2 in an RVC is better viewed as a full verb because: a) unlike English verb particles, it can be syntactically independent (e.g., *Guo4 ma3lu4* 'cross the street'); b) the path meaning encoded in V2 remains the same whether V2 appears in an RVC or in isolation; c) unlike some English words with Latinate origins (e.g., *ascend*, *descend*,

exit), the independent use of V2 in either oral or written medium incurs no sense of obsolescence.

Some other studies also suggest that Chinese seems more like an ‘equipollently-framed’ language (e.g., Chu, 2009; Gao, 2001; Chen and Guo, 2009; Ji *et al.*, 2011a). For instance, Ji *et al.* (2011b) revealed that the implications of being equipollently-framed can be evaluated from multiple perspectives. First of all, in terms of lexicalisation pattern, both manner and path of motion can be encoded in the verb (i.e., an RVC in Chinese). Secondly, if we go beyond the grammatical category of given elements, the syntactic strategy at the discourse level in Chinese looks quite similar to that of a typical verb-framed language, such as French: it encodes path of motion only in the verb whilst expressing manner in the periphery of an utterance via subordination or gerund (e.g., *Mao1tou2ying1 shan2 zhe chi4bang3 chu1-lai2 le.* ‘The owl, flapping its wings, exited [towards the speaker].’). Thirdly, to make the issue more complicated, elements like V2 in Chinese have undergone some significant changes in history. In Classical Chinese, these elements are undoubtedly full verbs (e.g. *Niu2yang2 xia4-lai2.* ‘The cows and sheep descended [towards us].’). These verbs are gradually weakened into prepositions in Modern Chinese, largely losing their verbal meanings. Seen in this way, the syntactic independence and the path meaning conveyed by V2 can be considered as a remnant of chronological change (see Peyraube 2006 for a detailed discussion).

As discussed above, we take Chinese as an ‘equipollently-framed’ language that is both similar (i.e., partly satellite-framed) and dissimilar (i.e., partly verb-framed) to the clearly satellite-framed English. Previous studies of language influences on spatial conceptualisation tend to focus on languages with opposing typological properties (satellite- vs. verb-framed). The proposed study, however, involves languages showing less dramatic differences (equipollently-framed Chinese and satellite-framed English), which will reveal whether effects of language are strong enough to lead to variations in mental representation of motion (if at all), even with minimal differences between languages.

4. Rationale of the study

The present study aims to investigate, in the particular domain of motion representation, the potential effect of language differences on motion event cognition in a triads matching task, using measurements of both overt choices as indices of

explicit processing and RT as reflections of implicit processing. Specifically, we aim to examine whether (and to what extent) the number of manner- (or path-) matches, as well as the RT to motion stimuli, varies significantly with language group and/or with age level. We generate two hypotheses regarding the research questions:

a. The ‘universal’ hypothesis: Talmy (2000) proposes a ‘Basic Motion Scheme’ in which path is considered the most central and indispensable semantic dimension for motion. A motion event, a translocational one in particular, is kept if manner of motion is not specified (please compare: He is *jumping* across the field vs. He is *going* across the field). However, without any path information, a motion event hardly exists, or at least will be transformed into one involving no changes of location (please compare: He is *jumping* across the field vs. He is jumping [in situ]). Seen in this way, it is likely, in certain circumstances, that typological differences between languages can be superficial and tend to be varied surface realisations of a common underlying conceptual framework. Given the cognitive salience of path-in-motion conceptualisation (Talmy, 2000), participants may evaluate the similarity between motion scenes on the basis of path-similarity, irrespective of language and age.

b. The linguistic relativity hypothesis: typological differences between languages should influence motion conceptualisation after the language-specific pattern for motion description has been fully acquired and internalised. We thus predict, in the first instance, that the effect of language on thought should be minimal at the onset of acquisition, and therefore, the two youngest groups of 3-year-olds would behave similarly in fixating more frequently on path-matched (vs. manner-matched) screens.

Older children of 8 years and adults, however, would show language-biased behavioural patterns. The crucial difference in motion event description between Chinese and English is the amount of information characteristically encoded in the verb. The verb domain is arguably the most marked grammatical category in a sentence, and any semantic information it carries is, thus, psychologically salient. In English, a single verb root encodes manner of motion only, whereas in Chinese, a verb compound packages manner, path and (optionally) deixis of motion simultaneously. Questions arise as to what motion event components standardly associate (or co-occur) with the mental representation of motion events. In English, it is the schema of ‘manner-salience’ that associates most frequently with motion event cognition, whereas, in Chinese, the schema of ‘manner-and-path salience’ appears concurrently with motion event representation. Such differences in the association

pattern may have important cognitive implications. According to the psychological theory of associative learning, (e.g. Bybee, 2006; Colunga and Smith, 2005), event representations “build up, or emerge over exposure to a number of specific instances of associations. The more routinized an association becomes, the easier it is to retrieve and utilise it for purposes of categorisation” (e.g. in tasks of similarity judgment, rearrangement of scenes, etc.; see Athanasopoulos *et al.* 2015: 141 for a detailed discussion). Seen this way, there is a high likelihood that English older children and adults will use ‘manner salience’ as a basis for their similarity judgment more frequently (i.e. more manner-matches), whereas their Chinese counterparts would rely on manner-salience and path-salience equally frequently in their choices (i.e. approximately equal number of manner- and path-matches).

We use two types of measures to assess the participant’s judgment of the similarity between motion scenes, that is, the overt choice of either manner- or path-match and the continuous variable of RT. It is reasoned that a forced choice (either A or B) in a timed judgment task would reflect primarily explicit processing whilst any differences in RT between manner- and path-matches would reflect online, implicit processing. The choice response data reveals what decision has been made, and at what frequency, whilst the RT data indicates the speed at which a given judgment is rendered.

The importance of including the RT data in our analysis can be seen from findings of previous studies suggesting that there is usually a discrepancy between overt choices and RT data in tasks such as grammaticality judgment and sentence interpretation. To cite an example, Kemp and MacWhinney (1999) use a picture-choice paradigm to examine possible language differences in online processing of morphological cues (e.g. word order, case marking, animacy, etc.) for agentivity. Russian and German participants listen to a transitive sentence while making a choice between two candidate agents that are presented visually; only one can be selected as the agent. It is reported that speakers of Russian and German behave similarly in many regards in their interpretation of a sentence. Both groups rely on case marking, the most reliable cue in both languages, to determine the agent of the sentence. However, the RT data reveals language differences: the magnitude of the processing efficiency from case marking is significantly larger in Russian, reflecting the fact that case marking is a stronger cue in Russian than in German (Kemp and MacWhinney 1999: 151).

Seen in this light, RT measurements can shed fresh insights into the decision-making process. It may reveal aspects of processing that are not readily available in results from choice response measures. In the current study, the typological difference between the two languages in comparison might not be sufficiently strong to have a significant impact on the final judgment, but it may still have a detectable effect on the speed with which participants react to the motion stimuli.

5. Methodology

5.1. Participants

Data was collected in universities, primary schools and nurseries in China and in UK, respectively. A total of 192 participants were recruited for this study (see Table 1, below, for details), and they were from three age levels (3 years, 8 years and adults) and two language groups (English and Chinese) in London and Beijing, respectively.

The age groups are chosen as representing critical points in a developmental pattern of spatial language acquisition attested from Ji *et al.* (2011a): age 3: mainly path expressions, unstable and atypical manner-path combinations; age 8: fully acquired typical manner-path use in native language.

Table 1

Groups of participants in the study

Group ID	Native language	Age level	Number of participants	Mean age
C03	Chinese	3 years	32	3;2 ($SD = 0.31$)
C08	Chinese	8 years	32	8;2 ($SD = 0.02$)
CAD	Chinese	Adults	32	19;3 ($SD = 0.97$)
E03	English	3 years	32	3;2 ($SD = 0.32$)
E08	English	8 years	32	8;3 ($SD = 0.03$)
EAD	English	Adults	32	21;0 ($SD = 0.17$)

5.2. Materials

The stimuli of the present study were 16 triads of short video clips showing voluntary motion events, each depicting a boy performing a self-instigated action showing

specific manners (e.g. *running*, *jumping*) and 4 types of trajectory: verticality (*up*, *down*), boundary-crossing (*across*, *into*), deixis (*towards*, *away from*) and trajectory encircling, or parallel to the ground of motion (*along*, *around*). In all stimuli, manner and path of motion were presented as occurring simultaneously and equally salient in perception. A complete list of the stimuli, along with an example illustration, is given in Appendices 1 and 2.

The 16 triads of stimuli comprised 48 voluntary motion events: 16 targets and 32 alternates (two for each of the target events). To give an example, the target video clip in triad No. 1 depicted a boy walking down the stairs; in its path-match alternate, only the manner of motion varied (the boy *jumped* down stairs), and in its manner-match alternate, only the path of motion was altered (the boy walked *up* stairs). All stimuli were organised into two randomised orders A and B, which were counterbalanced across participants within each language and each age group. The presentation position of manner-match vs. path-match video clips (i.e. left or right of the screen) was also counterbalanced across stimuli in a given order. Target and alternate videos appeared for 5 seconds each, which were then followed by 1 second of black screen.

5. 3. Procedures

The older children and adult participants were invited to watch video clips displayed on a MacBook Pro and asked to indicate their judgments of similarity between motion scenes by pressing certain keys on the keyboard. The stimuli were played through the stimulus presentation software ‘SuperLab 4.5’. This generated, at the end of the testing session, a file that included participant’s response type (manner-match or path-match) and his or her reaction time (RT) in milliseconds (ms).

The stimuli were played in a synchronised series, with the target videos playing first in the centre of a screen with a black background, followed by two simultaneous alternate videos placed side-by-side on the same screen. The participant listened to the audio instruction such as: ‘This is 1. Which one is most like 1?’ at the start of their viewing of a triad. A verbal interference task was utilised in which numbers in random sequence was ‘shadowed’ (repeated aloud) to older participants to prevent them from subconsciously verbalising motion events during judgment, thus eliciting truly ‘non-linguistic’ thinking.

For the 3-year olds, a ‘preferential looking’ scheme was adopted. The stimuli played via SuperLab on the laptop screen were projected onto a 16: 9 format

projection screen. A video camcorder was placed just above the screen in a central position to record the younger children's eye movements whilst they watched the video clips.

5. 4. Coding

The data was measured, in the first instance, by the categorical variable (i.e. preferences). The judgments of adults and older children were classified as either manner-match or path-match according to the button that was selected. In the preferential looking experiment, we coded children's eye fixation to manner-match and path-match screens. If children fixated more on the manner-match video than the path-match video, then they were considered more manner-oriented in similarity judgments.

The categorical analysis was supplemented by the continuous measurement of RT (in older children and adults) and fixation time (in the 3-year-olds). The RT for a given stimulus was calculated from the time between the presentation of alternate videos in a triad until their completion. Screening for outliers in the RT data was conducted by removing all observations that were more than two standard deviations from the group mean.

Video recordings of the younger children's eye movements were coded in a frame-by-frame mode using QuickTime Player 7.6. The measure used was the amount of time the child fixated on the manner- or path-matched screens. Twenty per cent of the video data (12 out of 64 files) were re-coded, and the *Kappa* index suggested that the reliability of the data reached ideal ($\kappa = 1.0$).

6. Results

This section aims to answer the following three questions: a. whether (and to what extent) the mean number of path-match (complementary with the number of manner-match) significantly varies with language (English, Chinese) and/or age (3 years, 8 years and adults; Sub-section 6.1); b. whether (and to what extent) the overall RT to motion event screens varies as a function of language, age (3 years, 8 years) and/or preference type (manner-match, path-match; Sub-section 6.2); c. whether (and to what extent) the 3-year-olds' fixation time to motion event screens varies with language and preference type (i.e. Sub-section 6.3). Depending on the specific question asked, statistical tests such as three-way mixed ANOVA were utilised to

explore relevant datasets.

6.1. Mean number of preferences for the path-match across 6 participant groups

The data of categorical preferences were represented as falling into one of the two major strategies in judging the similarity between screens: the manner-match and the path-match. For the youngest children, their preferences were determined by difference scores between the number of frames they fixated to the manner-match minus that to the path-match for each item. Thus, positive numbers represented a preference for the manner-match and negative numbers indicated a preference for the path-match. The mean was then calculated by recording the number matches out of 16 across individuals in a group.

Figure 1, below, presents the mean number of both manner-match and path-match preferences across 6 participant groups (C03, C08, CAD, E03, E08 and EAD).

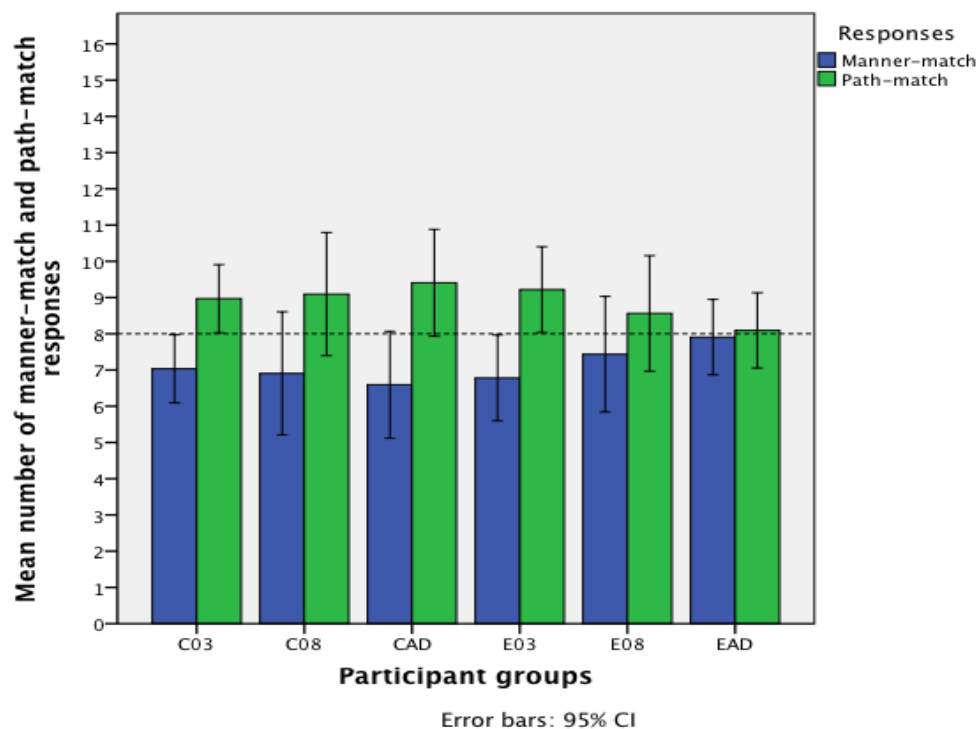


Figure 1. Mean number of manner-match and path-match preferences across participant groups

Given that the number of manner-match is complementary with that of path-match, and a visual inspection of the graph suggests a common tendency for the patch-match,

we conducted a two-way factorial ANOVA with language (English, Chinese) and age (3 years, 8 years and adults) as two between-subjects factors to test whether the mean number of the path-match preference varies with language and/or age. The analysis revealed no main effect of either language ($F(1, 192) = 0.963, ns$) or age ($F(2, 192) = 0.148, ns$), nor any significant interaction effect between the two variables ($F(2, 192) = 0.694, ns$).

Furthermore, in order to determine whether path-match is the predominating decision strategy across participant groups, a set of chance analyses was performed, separately, on each group, in which the mean number of path-matches was compared against a test value of 8 (i.e. the average of the group). The results confirmed that the number of path-match preferences fell above chance levels (more path-matches) for the two youngest groups (C03 and E03), but remained at chance level across all other groups (see Table 2).

Table 2

Statistical results of chance analyses conducted on the number of path-matches in each participant group

Group ID	Mean & SD of numbers path-match	95% confidence interval for mean	t value & significance (2-tailed)	Effect size (Cohen's d)
C03	$M = 8.969$ ($SD = 2.609$)	[8.023, 9.909]	$t(31) = 2.101$ $p = .044$	$d = .371$
C08	$M = 9.125$ ($SD = 4.654$)	[7.447, 10.803]	$t(31) = 1.367$ ns	N/A
CAD	$M = 9.406$ ($SD = 4.095$)	[7.930, 10.883]	$t(31) = 1.943$ ns	N/A
E03	$M = 9.219$ ($SD = 3.280$)	[8.036, 10.401]	$t(31) = 2.102$ $p = .044$	$d = .372$
E08	$M = 8.563$ ($SD = 4.421$)	[7.76, 11.11]	$t(31) = .720$ ns	N/A
EAD	$M = 8.094$ ($SD = 2.889$)	[7.052, 9.135]	$t(31) = 0.184$ ns	N/A

In order to test whether the mean number of path-matches varies significantly with individual items, an additional three-way mixed ANOVA was conducted with language (Chinese, English) and age (3 years, 8 years and adults) as two between-subjects factors, and test item (16 levels) as the within-subjects factor. The results indicated a significant effect of item, $F(15, 119) = 16.546, p < .001$, partial $\eta^2 = .082$, as well as an interaction between item and age, $F(30, 119) = 9.052, p < .001$, partial $\eta^2 = .089$. No other main effects or interaction between factors were detected. A closer examination of the data further revealed that item 2 (target: skateboarding out of house; alternates: skateboarding *along* house and *walking on stilts* out of house) elicited more path-matches in the group of 3-year-olds ($M = .766$) than in older children of 8 years ($M = 0.500$) and adults ($M = 0.437$), $p = .001$ in both cases. In contrast, item 11 (target: jumping along benches; alternates: jumping *down* benches and *waddling* along benches) was more path-inducing for the group of 8-year-olds ($M = .828$) and adults ($M = .906$) than for the 3-year-olds ($M = .188$), $p < .001$ in both cases. Such findings suggest that the conceptualisation of a particular item as more salient in manner or path varies significantly between children at the initial stage of linguistic and cognitive development and those who have (nearly) achieved linguistic and cognitive maturity.

6. 2. Older children and adults' RT in judgment

As revealed in the previous section, older children and adults in both English and Chinese did not show a particular preference for either manner-matched or path-matched screens in terms of overt choices. Thus, our second set of analysis tests whether there is any language/age difference reflected in the RT data. Overall RT (in ms) to the motion stimuli was analysed in two tests of three-way repeated measures ANOVA (i.e. by-participant and by-item analyses) with language (Chinese, English), age (8 years, adults) and preference type (manner-match, path-match) as factors. The results revealed main effects of language ($F_1(1, 124) = 18.929, p < .001$, partial $\eta^2 = .132$; $F_2(1, 15) = 103.349, p < .001$, partial $\eta^2 = .873$), main effects of age ($F_1(1, 124) = 34.512, p < .001$, partial $\eta^2 = .218$; $F_2(1, 15) = 163.572, p < .001$, partial $\eta^2 = .916$), as well as a main effect of preference type in the participants analysis ($F_1(1, 124) = 8.697, p = .004$, partial $\eta^2 = .066$; $F_2(1, 15) = 2.971, ns$). No significant interaction effects between variables were attested in both by-participant and by-item

analyses (i.e. Fig. 2; see also Table 3).

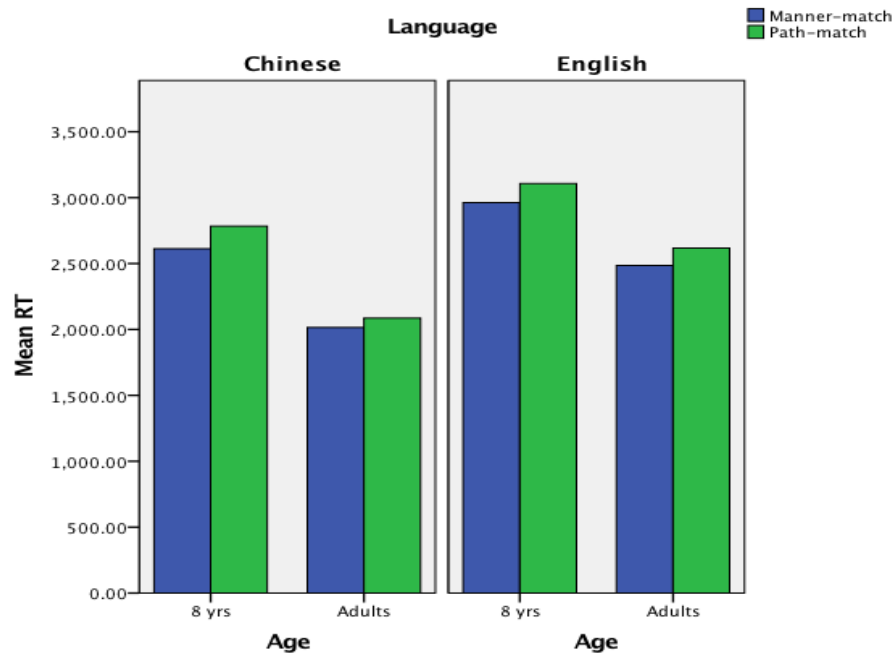


Figure 2. Mean RT (in ms) to motion stimuli as a function of language, age and preference type

Table 3

Mean RT (in ms) to motion stimuli across language groups, age levels and preference types

Language	Age	Preference	Means (and SDs)	95% Confidence Interval
Chinese	8 years (C08)	M-match	2612.582 (501.470)	[2397.244, 2827.919]
		P-match	2784.175 (494.019)	[2580.593, 2987.757]
	Adults (CAD)	M-match	2015.638 (568.159)	[1800.301, 2230.976]
		P-match	2085.913 (432.469)	[1882.331, 2289.495]
English	8 years (E08)	M-match	2962.956 (671.330)	[2747.619, 3178.293]
		P-match	3107.595 (655.623)	[2904.013, 3311.176]
	Adults (EAD)	M-match	2485.248 (700.080)	[2269.911, 2700.585]
		P-match	2618.052 (702.311)	[2414.470, 2821.633]

Generally, by age level, adults, irrespective of language group, were significantly quicker ($M = 2301.213$) in judging the similarity between motion screens than

children of 8 years ($M = 2866.827$, $p < .001$). In terms of preference type, participants across language and age groups used significantly shorter time in judging the manner-similarity ($M = 2519.106$) than in judging the path-similarity ($M = 2648.934$, $p = .004$). Importantly, by language group, Chinese older children and adults responded significantly quicker ($M = 2374.577$) to the motion stimuli as compared to their English counterparts ($M = 2793.463$, $p < .001$). A closer look at the data further revealed that they reacted quicker not only in manner-matched choices ($M = 2314.110$ in Chinese vs. $M = 2724.102$ in English, $p = .002$), but also in path-matched decisions ($M = 2435.044$ in Chinese vs. $M = 2862.823$ in English, $p = .001$).

6. 3. 3-year-olds' motion event looking time

As revealed earlier by the categorical data, the younger children of 3 years, irrespective of language, preferred path-match over manner-match. The same set of data was also measured in terms of fixation time (or potential degree of difference). A two-way mixed ANOVA, with language group (Chinese, English) as the between-subjects factor and preference type as the within-subjects factor, (manner-match, path-match) revealed a significant effect of preference type only, $F(1, 62) = 9.466$, $p = .003$, partial $\eta^2 = .132$. There was no interaction between language and preference type, $F(1, 62) = .281$, *ns*. Both Chinese and English 3-year-olds looked at path-matched screens for significantly longer than at manner-matched scenes (total mean of fixation to path-match = 2.468 sec vs. total mean of fixation to manner-match = 2.214 sec, $p < .001$). Thus, our analyses of younger children's behaviour, using categorical measure (manner- or path-matched choice) and continuous measure (fixation time), converged on a strong resemblance between Chinese and English 3-year-olds in being predominantly path-oriented.

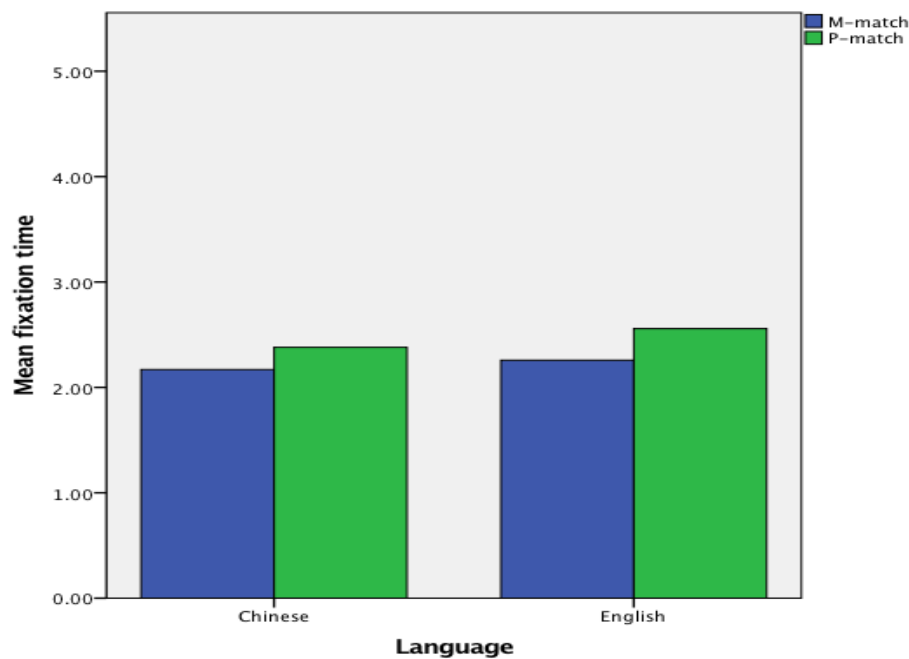


Figure 3. The 3-years-olds' fixation time to manner-matches and path-matches by language group (in sec)

Table 4

Means and Standard Deviations for fixation time to manner-match and path-match screens between Chinese and English 3-year-olds (in number of seconds)

Language	Fixation type	Mean fixation time	Std. Deviation
Chinese	M-match	2.169	0.359
	P-match	2.380	0.390
English	M-match	2.259	0.312
	P-match	2.557	0.332

7. Discussion and conclusion

In this paper, we are generally interested in whether, and to what extent, language-specific influences can go beyond the level of language use into the realm of cognition. We specifically looked at how English and Chinese speakers of different ages judged the similarity between motion events. To this end, we analysed not only overt choices on the part of speakers (i.e. manner- or path-match), but also behavioural responses made by speakers during the testing session (i.e. RT and eye fixation).

Our results showed, first of all, that the Chinese and English 3-year-olds shared a

tendency to be more path-oriented, which was evidenced by their significantly longer fixation on path-matched (vs. manner-matched) screens. Thus, young children's conceptualisation of motion was similar prior to becoming accustomed to using motion expressions in ways typical to their native language. Such a finding confirms Talmy's (2000) 'Basic Motion Scheme' in which path (rather than manner) is a cognitively predominant dimension in the motion conceptualisation of young children.

For groups of older children (i.e. the 8-year-olds) and adults, in the first instance, our results suggested that participants performed just at the chance level in their similarity judgments, irrespective of language and age group. The analyses of the RT data, however, revealed significant differences in three aspects, viz., language (i.e. Chinese participants were generally quicker than their English peers in decision-making), age (i.e. adults were significantly quicker than older children of 8 years) and type of choice (i.e. manner-matches were decided in significantly shorter time than path-matches). Such a divergence of results highlights the importance of including the RT measure in our analysis: it reveals aspects of behavioural responses that cannot be readily attested in the forced choice data.

The first question arising from our findings concerns overt choices. The absence of a language-biased behavioural pattern between Chinese and English speakers does not come as a surprise. Some previous studies of motion event cognition already suggest no (or unclear) variations in behavioural response, even between speakers of languages with opposing typological status (e.g. satellite-framed English vs. verb-framed Greek in Papafragou *et al.* [2002]). Several perspectives might be taken on this issue. The typological distance between Chinese (equipollently-framed) and English (satellite-framed) might not be sufficiently great to allow any differences in categorisation to surface. Or, the measure of choice response might not be sufficiently sensitive to language-engendered differences in behaviour.

A more plausible explanation of these results, however, might be related to a key question in the field of motion event typology, that is, the nature of typological contrast between languages in motion description. In fairness, the bipartite system as proposed by Talmy (i.e. satellite- vs. verb-framed) is not intended to be a clear-cut classification: "these typological characterisations often reflect tendencies rather than absolute differences between languages" (Berman and Slobin 1994: 118). The typology *per se* focuses on the most "characteristic" way of encoding a motion event,

that is, a pattern which is colloquial in style, frequent in occurrence and pervasive in scope. By this is meant that more than one type of framing patterns can actually be observed in a given language. Recent investigations reveal that some typical verb-framed languages, such as French and Italian, seem to have clear satellite-framed features whilst some traditional satellite-framed languages (e.g. Tzeltal and Chinese) have path-expressing verbs. As a case in point, Iacobini and Masini (2006) investigate verb-particle constructions and prefixed verbs in Italian and find that Italian employs both the ‘Romance type’ and the ‘Germanic type’ of expression of motion events. More specifically, path information is conflated with motion in verb roots, but it is not the only, or the privileged, way of realising the path feature in Italian. Rather, the primary function of Italian post-verbal particles is to add directional values to the verbal root and they function as true satellites, just as in Germanic languages (2006: 161). Similarly, even in English, a language confirmed to be predominantly satellite-framed, there is a wealth of path verbs, including not only less frequently used Latinate borrowings such as *ascend*, *descend*, *enter* and *exit*, but also more frequently used colloquial ones such as *rise*, *fall* and *sink* (see, for instance, Beavers *et al.*, 2010; Brown, 2004; Kopecka, 2006; Iacobini and Masini, 2006; Talmy, 2009). The most important implication of great within-language variability is that the widely attested crosslinguistic differences in motion description might be more probabilistic than categorical. It thus limits the extent to which speakers of different languages construe motion events in categorically different ways and further restricts any language-specific behavioural patterns in nonverbal tasks (see Montero-Melis 2016: 642 for a detailed discussion).

Secondly, our RT data analysis indicates a significant difference between preference types. Despite the fact that older children and adults rely on manner- and path-similarity criteria equally frequently in their judgments, they actually react quicker in making manner-match choices than in opting for path-match decisions. Such an observation can be interpreted in different ways. One possibility is that manner is a more straightforward and predictable dimension than path. The agent’s manner of motion does not alter throughout the movement and it can be perceived and compared more quickly in a few initial frames of the video clips. In contrast, one needs to wait to see a substantial portion of path in order to decide whether it is similar to the target path. Additionally, the age-related differences in RT can be seen from the perspective of cognitive development. The speed of processing is found to

increase with age and the decision-making becomes highly efficient in adulthood. Such changes with age are believed to be closely associated with development in working memory and various aspects of thought (e.g. Demetriou *et al.* 2002).

Importantly, our results reveal that the language differences in RT tend to be statistically significant. Chinese participants, on the whole, respond more quickly to motion stimuli than their English counterparts. In fact, they use significantly shorter time in both manner- and path-matched choices. Such findings can be approached from varied perspectives. This phenomenon may be related to different types of decision strategies between the two groups of speakers (e.g. a ‘jump-to-conclusion’ strategy on the part of Chinese speakers vs. a ‘wait-and-see’ attitude on the part of English participants). It may also be related to culturally specific thought patterns. Previous studies suggest that East Asians (e.g. Chinese and Japanese) tend to have a ‘holistic’ thought pattern whereas Westerners usually have an ‘analytical’ way of thinking. The holistic process of judging objects and events can put East Asians at an advantage in behavioural tasks. For example, in Masuda and Nisbett’s study (2001), the Japanese speakers are found to be overall faster in RT than the American participants in recall and recognition tasks, due to their ability to incorporate contextual information into their construal of events and to view subparts of a particular scene as an interrelated and indecomposable whole.

There is also a likelihood in which the difference in processing speed is an outcome of different modes of information processing. Previous studies of high-level categorisation of natural images reveal that subparts of complex natural scenes can be processed in parallel very rapidly (see, for instance, Rousselet *et al.* [2002]). Questions thus arise pertaining to our specific task: what motion components are retrieved from short-term memory and processed in which fashion by the participant? As discussed earlier, path is the cognitively salient component for motion event conceptualisation and manner has a high degree of ‘codability’ (Slobin 1996) in both Chinese and English. Further, these two ingredients for motion are presented as occurring simultaneously in the experimental design. It is thus reasonable to believe that both manner and path of motion are retrieved for the purpose of categorisation. In other words, both Chinese and English participants in our task have the same potential for processing distinct motion components in parallel for event construal. But whether this potential can be realised and reflected in participants’ habitual behaviour (i.e. as measured by RT) may largely depend on the specific properties of a given language.

The structure of Chinese may prompt its speakers to realise this potential more than the structure of English. Specifically, the lexicalisation pattern for motion expressions in Chinese (i.e. both manner and path are equally prominent as encoded in compound verbs) encourages its speakers to evaluate the manner and path dimensions simultaneously for final judgment (i.e. the mode of parallel processing), whereas the motion lexicalisation pattern in English (manner is more prominent in the verb root than path in periphery) biases its speakers to assess the manner similarity in the first instance, which is immediately followed by an evaluation of the path dimension. A final decision can only be made after the salience of the two dimensions has been weighed. Such a serial processing may not be strictly sequential in our particular experimental context; overlaps of the successive processing times on distinct motion components may exist. It is speculated that Chinese participants process motion components in a highly efficient ‘parallel’ way, whereas their English counterparts process the same set of components in a primarily ‘sequential’ way, which may have slowed down their reaction time in judgments.

Overall, our findings are not entirely consistent with any hypothesis postulated in Section 4. The complexity of motion stimuli in the task seems to have a major impact on the results. We investigate, in another study, how Chinese and English speakers express caused motion events such as *The boy kicked the ball up the hill* [*The ball rolled up to the hill and the boy walked with the ball*] (Ji and Hohenstein, under review). These caused motion stimuli represent complex events with multiple motion components. To illustrate, in the aforementioned example, apart from cause of motion (as encoded in *kick*) and path of motion (as encoded in *up*), three specific types of manner information are presented, viz., manner of cause (*kicking*), manner of the object (*rolling*) and manner of the agent (*walking*).

The results of the caused motion study suggest, first of all, that the two younger groups of 3-year-olds are predominantly path-oriented (similar to the findings in the current study). Using categorical measurement of overt choices, older children and adults also show a shared tendency of being more path-oriented. However, the analysis of the RT data reveals significant variations in motion event cognition that can be related to linguistic differences. English older children and adults are found to react significantly more quickly to manner-matched screens than to path-matched ones, whereas their Chinese counterparts tend to spend an approximately equal amount of time in both manner- and path-matched judgments. This divergence can

arguably map onto the typological status of relevant languages in motion event typology (i.e. satellite-framed English vs. equipollently-framed Chinese). Unlike the findings in the present study, the set of results regarding the caused motion stimuli indicates a likelihood that children's non-linguistic thought is similar prior to internalising the lexicalisation pattern of motion events in their native languages, but shows divergences after such habitual use.

Taken together, our research seems to highlight the complexity of the relationship between language, space and mind. Various factors need to be weighed at the same time in assessing the effect of language differences on motion event cognition, for instance, the typological distance between languages (satellite-, verb-, and equipollently-framed), the nature of motion events under investigation (voluntary vs. caused), the type of measurement used (categorical vs. continuous) and the design features of test items (more manner- vs. path-inducing).

References

- Allen, S., Ozyurek, A., Kita, S., 2007. Language-specific and universal influences in children's syntactic packaging of Manner and Path: A comparison of English, Japanese, and Turkish. *Cognition* 102(1), 16–48.
- Athanasopoulos, P., Bylund, E., Monteromelis, G., Damjanovic, L., Schartner, A., & Kibbe, A. 2015a. Two languages, two minds: flexible cognitive processing driven by language of operation. *Psychological Science* 26(4), 518.
- Beavers, J., Levin, B., & Tham, S. W. 2010. The typology of motion expressions revisited. *Journal of Linguistics* 46(2), 331–377.
- Berman, R., & Slobin, D. I., 1994. *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Berthele, R., 2013. Disentangling manner from path: evidence from varieties of German and Romance. In: Goschler, J., Stefanowitsch, A. (Eds.), *Variation and Change in the Encoding of Motion Events*. John Benjamins, Amsterdam, pp. 55–76.
- Boroditsky, L., 2000. Metaphoric structuring: Understanding time through spatial metaphors. *Cognition* 75(1), 1–28.

- Bowerman, M., 1999. Learning how to structure space for language: A crosslinguistic perspective. In: Bloom, P. M., Peterson, L. (Eds.), *Language and Space*. MIT Press, Cambridge, MA, pp. 385–436.
- Bowerman, M., Choi, S., 2003. Space under construction: Language-specific categorization in first language acquisition. In: Gentner, D., Goldin-Meadow, S. (Eds.), *Language in Mind: Advances in the Study of Language and Thought*. MIT Press, Cambridge, MA, pp. 387–427.
- Brown, P. 2004. Position and motion in Tzeltal frog stories: The acquisition of narrative style. In Stromqvist, S., Verhoeven, L. (Eds.), *Relating Events in Narrative: Crosslinguistic and Cross-contextual Perspectives* (37-58). Earlbaum, Mahwah, N. J., pp. 37–58.
- Bybee, J. 2006. *Frequency of Use and the Organization of Language*. Oxford University Press, Oxford.
- Chen, L., Guo, J., 2009. Motion events in Chinese novels: evidence for an equivalently-framed language. *Journal of Pragmatics* 41, 1749–1766.
- Choi, S., Bowerman, M., 1991. Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41, 83–121.
- Chu, C. Z., 2009. Path of motion: Conceptual structure and representation in Chinese. In: Xing, Z. (Ed.), *Studies of Chinese Linguistics: Functional Approaches*. Hong Kong University Press, Hong Kong, pp. 65–73.
- Colunga, E., & Smith, L. B. 2005. From the lexicon to expectations about kinds: a role for associative learning. *Psychological Review* 112(2), 347–382.
- Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. 2002. The development of mental processing: Efficiency, working memory, and thinking. *Monographs of the Society of Research in Child Development*, 67, Serial Number 268.
- Dittmar, M., Abbot-Smith, K., Lieven, E., Tomasello, M., 2008. Young German children's early syntactic competence: a preferential looking study. *Developmental Science* 11(4), 575–582.
- Filipovic, L., Jaszczolt, K. (Eds.), 2012. *Space and Time in Languages and Cultures: Linguistic Diversity*. John Benjamins Publishing Company, London.
- Flecken, M., 2011. Event conceptualization by early Dutch-German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition* 14(1), 61–77.

- Gao, H., 2001. Notions of motion and contact for physical contact verbs. In: Holmer, A., Svantesson, J., Viberg, A. (Eds.), *Proceedings of the 18th Scandinavian Conference of Linguistics*, vol. 2. University of Lund, Lund, pp. 193–209.
- Gennari, S., Sloman, S., Malt, B., Fitch, W., 2002. Motion events in language and cognition. *Cognition* 83, 49–79.
- Hickmann, M., 2006. The relativity of motion in first language acquisition. In: Hickmann, M., Robert, S. (Eds.), *Space in Languages: Linguistic Systems and Cognitive Categories*. John Benjamins Publishing Company, Amsterdam, pp. 281–308.
- Hickmann, M., Hendriks, H., 2010. Typological constraints on the acquisition of spatial language in French and English. In: Hendriks, H., Hickmann, M., Lindner, K. (Eds.), *Space, Language and Cognition: Development Perspectives* Special issue of *Cognitive Linguistics* 21(2), pp. 181–188.
- Hohenstein, J., 2005. Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognition and Development* 6, 403–425.
- Iacobini, C. & Masini, F. 2006. The emergence of verb-particle constructions in Italian locative and actional meanings. *Morphology* 16, 155–188.
- Ji, Y., Hendriks, H., Hickmann, M., 2011a. How children express caused motion events in Chinese and English: Universal and language-specific influences. *Lingua* 121, 1796–1819.
- Ji, Y., Hendriks, H., Hickmann, M., 2011b. The expression of caused motion events in Chinese and in English: some typological issues. *Linguistics* 49(5), 1041–1076.
- Ji, Y., & Hohenstein, J. (under review). English and Chinese children's motion event similarity judgments. *Cognitive Linguistics*.
- Jackendoff, R., 1996. The architecture of the linguistic-spatial interface. In: Bloom, P., Peterson, M., Nadel, L., Garrett, M. (Eds.), *Language and Space*. MIT Press, Cambridge, MA, pp. 1–30.
- Kempe, V & MacWhinney, B. 1999. Processing of morphological and semantic cues in Russian and German. *Language, Cognition and Neuroscience*, 14(2), 129–171.
- Kopecka, A. (2006). The semantic structure of motion verbs in French: Typological perspectives. In Hickmann, M., Robert, S. (Eds.), *Space in Languages: Linguistic Systems and Cognitive Categories*. John Benjamins, Amsterdam, pp. 83–101.

- Landau, B., Lakusta, L., 2006. Spatial language and spatial representation: Autonomy and interaction. In: Hickmann, M., Roberts, S. (Eds.), *Space in Languages: Linguistic Systems and Cognitive Categories*. John Benjamins Publishing Company, Amsterdam, pp. 309–333.
- Levinson, S., 2003. *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge University Press, Cambridge.
- Lucy, J., 1992. *Grammatical Categories and Cognition*. Cambridge University Press, Cambridge.
- Malt, B., Sloman, S., Gennari, S., 2003. Universality and language specificity in object naming. *Journal of Memory and Language* 49(1), 20–42.
- Masuda, T., & Nisbett, R. E. 2001. Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *Journal of Personality & Social Psychology*, 81(5), 922–934.
- McWhorter, J. H., 2014. *The Language Hoax: Why the World Looks the Same in Any Language*. Oxford University Press, Oxford.
- Montero-Melis, G., Jaeger, T. F., & Bylund, E. 2016. Thinking is modulated by recent linguistic experience: second language priming affects perceived event similarity. *Language Learning*, 66(3), 636–665.
- Papafragou, A., Massey, C., Gleitman, L., 2002. Shake, rattle, ‘n’ roll: The representation of motion in language and cognition. *Cognition* 84, 189–219.
- Papafragou, A., Hulbert, J., Trueswell, J., 2008. Does language guide event perception? Evidence from eye movements. *Cognition* 108(1), 155–184.
- Peyraube, A. 2006. Motion events in Chinese: A diachronic study of directional complements. In: Hickmann, M., Roberts, S. (Eds.), *Space in Languages: Linguistic Systems and Cognitive Categories*. John Benjamins, Amsterdam, pp.121–135.
- Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. 2010. Parallel processing in high-level visual scene categorization. *Nature Neuroscience*, 5(7), 629–630.
- Slobin, D. I., 1996. Two ways to travel: Verbs of motion in English and Spanish. In: Shibatani, M., Thompson, S. A. (Eds.), *Grammatical Constructions: Their Form and Meaning*. Oxford: University Press, Oxford, pp.195–217.
- Slobin, D. I., 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In: Strömquist, S., Verhoeven, L. (Eds.), *Relating Events in Narrative Vol. 2: Typological and Contextual Perspectives*. Lawrence

- Erlbaum Associates, Mahwah, NJ, pp. 219–157.
- Soroli, E., Hickmann, M., 2010. Language and spatial representations in French and in English: evidence from eye-movements. In: Marotta, G., Lenci, A., Meini, L., Rovai, F. (Eds.), *Space in Language*. Editrice Testi Scientifici, Pisa, pp. 581–597.
- Talmy, L., 1985. Lexicalization patterns: semantic structure in lexical form. In: Shopen, T. (Ed.), *Language Typology and Syntactic Description, Vol. 3*. Cambridge University Press, Cambridge, pp. 36–149.
- Talmy, L., 2000. *Toward a Cognitive Semantics, Vol. 2: Typology and Process in Concept Structuring*. MIT Press, Cambridge, MA.
- Talmy, L., 2009. Main verb properties and equipollent framing. In: Guo, J. *et al.* (Eds.), *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin*. Psychology Press, New York, pp. 389–402.
- Von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., Schmiedtová, B., 2012. How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics* 50(4), 833–867.
- Zlatev, J., Blomberg, J., 2015. Language may indeed influence thought. *Frontiers in Psychology* 6, 1631.

Appendix A 16 voluntary motion items in the experiment

Item	Target	Manner-match	Path-match
1	Saunter away from fountain	Saunter towards fountain	Hop away from fountain
2	Skateboard out of house	Skateboard along house	Walk on stilts out of house
3	Run away from slope	Run down slope	Jump away from slope
4	Hop along trees	Hop towards trees	Jog along trees
5	Walk towards house	Walk out of house	Jog towards house
6	Crawl into cave	Crawl up cave	Jump into cave
7	Skate away from sculpture	Skate around sculpture	Sledge away from sculpture
8	Hop out of room	Hop into room	Limp out of room
9	Jump around flower stands	Jump along flower stands	Crawl around flower stands
10	Skip down bank	Skip along bank	Slide down bank
11	Jump along benches	Jump down benches	Waddle along benches
12	Drive car uphill	Drive car towards hill	Ride horse uphill
13	Walk on stilts around playground	Walk on stilts across playground	Skateboard around playground
14	Tiptoe down barn	Tiptoe across barn	Hop down barn
15	Run into post office	Run away from post office	Limp into post office
16	Walk down stairs	Walk up stairs	Jump down stairs

Appendix B Illustrations of the video stimuli showing voluntary motion events



a. The boy walking down stairs



b. The boy walking *up* stairs

c. The boy jumping down stairs

Audio stimuli accompanying the video were:

Target: This is one.

Alternates: Which one is most like one?